Research for the Credit Card Default prediction model via machine learning

* Prepare the training set, a records that are already have known class label.
* Build the model by applying one of learning algorithm using training set.
* Applied the model upon unknown data test set class.
* Evaluate the accuracy of the model.

# Paper 1

<http://www.ibii-us.org/Journals/JMSBI/V2N2/Publish/V2N2_6.pdf>

The comparison of machine learning methods to achieve most cost-effective prediction for credit card default

Purpose of the paper is that it is different from other papers with regards to machine learning insofar that it takes into account that the best net to check if customers are defaulting includes a mix of risky and not risky customers, as you could be turning away a lot of good customers.

# Paper 2

<https://www.researchgate.net/profile/Haifeng_Wang38/publication/319689046_Real_Time_Credit_Card_Default_Classification_Using_Adaptive_Boosting-Based_Online_Learning_Algorithm/links/59b991e1458515bb9c48a3f8/Real-Time-Credit-Card-Default-Classification-Using-Adaptive-Boosting-Based-Online-Learning-Algorithm.pdf>

Real Time Credit Card Default Classification Using Adaptive Boosting-Based Online Learning Algorithm

Adaboost: combine many weak classifiers into a single strong classifier

# Paper 3

<https://machinelearningmastery.com/boosting-and-adaboost-for-machine-learning/>

Learning about boost and adaboost

# Paper 4

<http://www.ijcte.org/papers/377-G1104.pdf>

Comparative Evaluation of Predictive Modeling Techniques on Credit Card Data

# Paper 5

<http://aircconline.com/mlaij/V3N1/3116mlaij01.pdf>

DEVELOPING PREDICTION MODEL OF LOAN RISK IN BANKS USING DATA MINING

3 classification algorithms

J48, bayesNet and naiveBayes

J48 split in trees, using the c4.5 algorithm

<https://www.geeksforgeeks.org/decision-tree-implementation-python/> (just decision tree)

<http://art.uniroma2.it/basili/MLWM09/002_DecTree_Weka.pdf>

BayesNET

naivveBayes

widely used, simple goood

J48 got the best result, high accuracy and low mean absolute error